Robert_S3_L09

Thu, 2/17 4:01PM 12:54

SUMMARY KEYWORDS

least squares method, simple linear regression, sample standard deviation, beta, coefficient, correlation, equation, predict, forecasting, called, y values, least squares, variable, y hat, regression, talk, slope, formula, intercept, causation

SPEAKERS

Robert McKeown

R Robert McKeown 00:05

Hello, and welcome to our first video on simple linear regression. What I want you to think about when it comes to simple linear regression is it's a method for making predictions. It's a tool for forecasting. We'll talk about the conditions under which it's a good tool to use for forecasting, and a few minutes. But I want you to think about that and have that in your head. Simple linear regression can be used for forecasting, it can be used in a business setting, you can be business can be forecasting, inventory, or sales. You can use it in economics, to forecast macroeconomic variables. And you can use it in other disciplines in the social sciences as well. Now economists often use regression to try to explore the nature of causation. So you learn that correlation is not causation, we're not going to talk about those more sophisticated regression techniques that try to squeeze out causation. Rather, I want you to think about simple linear regression as using X, some X variable to predict Y. And we're going to walk through the process by which we do this with something called least squares. Simple linear regression uses a linear model, that is a Y variable is equal to the slope M, multiplied by an X value, plus the Y axis intercept. Now in statistics, it's common to rearrange this linear equation that we see in front of us on the slides as follows, we're going to let the this be or beta be beta naught or beta underscore zero. And we're going to let the slope be beta one. And we do that because sometimes when we do a technique called multiple regression, we might have more than 1x. Variable. But we're not going to talk about that, in this module, we're just going to focus on simple linear regression, that's going to teach you how you can use regression to forecast is called simple, because there is just 1x variable. And that's where this simple comes from. Let's return our attention to the equation that we have here. And there's one thing about it that haven't explained yet that might strike you as unusual. And that's this Y hat, right here. Why does the Y have this little tilde? Or? Or I should say, Y circumflex. But we call it Y hat. Why do we have this new feature over the Y? Well, it's because this y is not the y that we see in the actual data, it's not the y value from the series of Y values that we have. It is the predicted value of Y. And if we want it to be use a probability about it, we could say it is the expectation of Y conditional on X. So Y hat is the predicted value of Y, these Y being the Y values and the data. Beta naught is the intercept as I explained, and beta one is the slope. These are known as the coefficients. And these coefficients, we're going to estimate them, and we're going to estimate them using something called least squares. How does it work? While the least squares method, least squares method uses the X and Y values in their series. So the observations that we get or the data that we have, and we're going to use that data to compute these coefficients. And then we're going to generate what's called the regression line, or the line of best fit, and also known as the Line of Least Squares. It's got many names, but all those lines are the same thing. What is this least squares method? What are these regression lines and these line of best fit is going to look like? Well I've drawn a little diagram here on the slide. We've got X and Y axis. So we've got some X values, and we've got some Y values and I'm going to draw them accordingly, like so. And the least squares method is going to draw a regression line, or a line of best fit through these observations, in order to minimize the distance between each observation and the line itself. We'll talk more in detail about how we do that. But essentially, we want to put a line through this cloud, this cloud of observations or this group of observations, that has the best fit. Before we continue on with some computation, I'll just quickly mention a little terminology, any economics, and broadly in the social sciences, these, and in business, these Y and X values, here, they have many names, my preferred name is to call the Y hat or the Y variable, the dependent variable, because its value depends on the independent variable X. So x is independent and can take on any value, Y depends on X, because Y, or at least Y hat is going to follow this equation right here. The least squares method uses just the standard deviations of X and Y, and the correlation between X and Y to compute the slope coefficient, beta one. So in order to use the least squares method, we're going to use this equation this formula right here. And if you take a more advanced class, in statistics or econometrics, you'll be able to derive this formula yourself here. Today, I'm just going to give you the formula for calculating this beta one coefficient. Now let's turn our attention back to the beta one equation down here, we can see that this slope coefficient beta one is closely related to the correlation. In fact, it's equal to the correlation multiplied by the ratio of the sample standard deviation divided up Y divided by the sample standard deviation of x. So we have r is the correlation between X and Y. S subscript y is the sample standard deviation Y. And S subscript x is a sample standard deviation of x. What is there some intuition that we can look at here? There is beta one is equal to the correlation between X and Y, multiplied by the ratio of Y to X, specifically, the variability in Y relative to the variability in X. Why is this important? Well, remember that coral correlation has no units. And it's always between oops, not zero, that's not right, negative one and one. While Sy is in units, of y. And Sx is in units of X. When we divide Sy by Sx, we get the ratio of variability of Y per unit of variability of X. And so in some sense, it's kind of like, you're going to get a ratio, you're going to get a ratio there, and it's going to kind of cancel out the units. And it's also going to help you it's also going to let the equation convert the X units into the Y units that we want to predict. Now there is a alternative equation for beta one I could rewrite beta one, replacing the correlation with its definition and it's definition is we have the sample covariance divided by the sample standard deviation of x times the sample standard deviation of y. There's, that's our definition of the correlation coefficient multiplied by the sample standard deviation of y over the sample standard deviation of x. And we can see that the sample standard deviation of y cancels out. And we're just going to be left with beta one is equal to the covariance between X and Y divided by the sample standard deviation of x. So we can see that this slope term and our simple linear regression is going to be directly related to the relationship between the y values and X bias, and it's going to show up either using the correlation coefficient or using the COVID covariance. So here, we have our equation, our formula for finding the beta one coefficient. Now let's turn our attention to finding our intercept or beta naught coefficient. The least squares method is going to calculate this for us using the means the mean of X and the mean of Y. And the formula we're going to use is beta naught is equal to the average value, the mean of Y minus the beta one coefficient that we calculated previously, multiplied by the mean value of X, the intercept is going to be equal to the difference between these two terms. Remember, Y bar is the mean of Y. And X bar is the mean of X. Now you have all the tools that you need to do some forecasting, you can predict Y given an X value, and you're going to do that by calculating these beta coefficients. And then you can plug in any X value that you want, and come up with a predicted Y value. And we're going to do that in the next video with an example from professional sports.